

MENU**SEARCH****INDEX****JAPANESE****BACK**

5 / 5

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 04-248642

(43)Date of publication of
application : 04.09.1992

(51)Int.Cl. G06F 12/06
G06F 11/10
G06F 12/16
G06F 15/347
G11C 29/00

(21)Application number :	03-204045	(71) Applicant :	IOBST KENNETH W RESNICK DAVID R WALLGREN KENNETH R
(22)Date of filing :	18.07.1991	(72)Inventor :	IOBST KENNETH W RESNICK DAVID R WALLGREN KENNETH R

(30)Priority

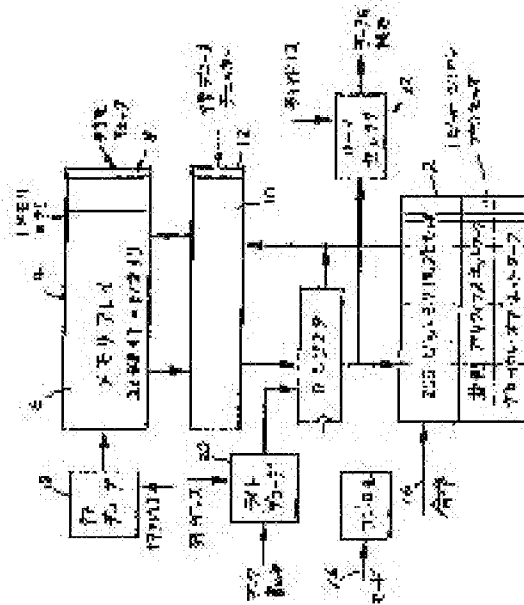
Priority number :	91 643633	Priority date :	18.01.1991	Priority country :	US
----------------------	--------------	--------------------	------------	-----------------------	----

(54) PIM CHIP FOR MEMORY INTEGRATED CIRCUIT AND ITS CONTROLLING METHOD

(57) Abstract:

PURPOSE: To maximize instruction/data bandwidth between a processor and a memory by mutually connecting the memory and a calculation function on the same integrated circuit and almost removing I/O terminals between data pins.

CONSTITUTION: A reference constitution element in a process-in memory circuit is a bit serial processor 2 provided with an additional local memory element 4. The element 4 can move one bit to/from the processor 2 through error correction circuit logic in each clock cycle. Processing in a memory chip is designed so as to utilize band width generated by the connection of a memory part to an operation part on the same integrated circuit. Since the chip includes many single bit serial processors 2 to be driven in parallel, the calculation of single instruction multiple data can be also quickly processed, and since error correction logic is also integrated in the chip, the generation of an error in memory data can be detected and corrected.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against
examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平4-248642

(43) 公開日 平成4年(1992)9月4日

(51) Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 12/06	5 3 0	8841-5B		
11/10	3 3 0 D	7832-5B		
	K	7832-5B		
12/16	3 2 0 F	7629-5B		
15/347	F	6798-5L		

審査請求 未請求 請求項の数18(全 11 頁) 最終頁に続く

(21) 出願番号	特願平3-204045	(71) 出願人	591176960 ケネス ダヴリユー、 アイオブスト KENNETH W. IOBST アメリカ合衆国 メリーランド 20904 シルバースプリング ウインドミル コー ト 18
(22) 出願日	平成3年(1991)7月18日	(71) 出願人	591176971 デビッド アール、 レスニック DAVID R. RESNICK アメリカ合衆国 ウィスコンシン 54703 オウ クレア ブラック アベニュー 2006
(31) 優先権主張番号	6 4 3 6 3 3	(74) 代理人	弁理士 蓼 経夫 (外2名)
(32) 優先日	1991年1月18日		最終頁に続く
(33) 優先権主張国	米国 (U S)		

(54) 【発明の名称】 メモリ集積回路のP I Mチップおよびその制御方法

(57) 【要約】

【目的】 データピン間の入出力端子をほとんどをなくすことによってプロセッサとメモリとの間の命令/データ帯域幅を最大にするためメモリ部と演算部を同一の集積回路上に結合したP I Mチップを提供すること、およびアドレス及びデータバスを介してP I Mチップのメモリバンクを制御する。

【構成】 メモリチップ内の処理は、同一集積回路上でメモリと計算とを結合させて、結合によって生じる帯域幅を利用できるように設計されている。チップには、並行に駆動される多数の単一ビット計算プロセッサが含まれている。エラー訂正論理もチップに組み込まれて、メモリデータ内のエラーを発生時に検出して訂正できるようにしている。

1

【特許請求の範囲】

【請求項1】 メモリ及び他のプロセッサからのデータを処理する装置であって、

(a) それぞれ複数のデータ入力部、複数の第1コマンド入力部(F1- μ)及び1つの出力部を備えており、またそれぞれが各クロックパルスで前記コマンド入力部からのコマンドに従って新しいデータ及び先に計算されているデータのうちの一方を選択する3つのセクタ装置(A, B, C)と、

(b) 前記セクタ装置の出力部に接続されて、前記セクタ装置からのデータの算術及び論理機能を計算して出力を発生する機能ユニットと、

(c) 機能ユニットの出力をセクタ装置入力部、別のプロセッサ及びメモリのうちの1つへ選択的に送る手段と、

(d) 新しく計算されたデータ及び基準データを外部コマンドに従って選択的に記憶する手段とを有していることを特徴とする装置。

【請求項2】 前記セクタ装置がマルチプレクサであることを特徴とする請求項1の装置。

【請求項3】 前記データ入力部は、メモリから新しいデータを、他のプロセッサから新しいデータを、前記機能ユニット出力部から送り戻されてきた先に計算されているデータを、さらにコマンド制御に従った強制データ出力を受け取ることを特徴とする請求項2の装置。

【請求項4】 前記機能ユニットは算術論理ユニットを有していることを特徴とする請求項3の装置。

【請求項5】 前記算術論理ユニットは、合計、繰り上げ及びストリング比較機能を実施することを特徴とする請求項4の装置。

【請求項6】 任意の関係オペレータ間で汎用並列プリフィクス演算機能を実施するためのネットワーク装置であって、

(a) ゼロレベルから始まる複数の連続的な通信レベルを定める手段と、

(b) Lがレベル数を表す時に各々が2のL乗のプロセッサを有する複数の連続グループを前記レベルの各々に定める手段とを有しており、

(c) 1グループ内の各プロセッサには、先行グループからの出力を有する単一の入力に対応しており、これによって一連の命令がレベルゼロからレベルLまでのレベルに対応して発行されて、2のL乗の値である並列プリフィクスが計算されるようになっており、

(d) レベルL及び後続レベルの入力が、先行入力のすべてを受け取っている各グループの単一のプロセッサに対応していることを特徴とする装置。

【請求項7】 あるレベルの前記グループは連続対に配列されており、各対の一方のグループがデータを前記対の他のグループへ送ることによって、並列プリフィクスの演算演算を定めることを特徴とする請求項6の装置。

2

【請求項8】 あるレベルのグループのうちの最後のグループからの出力が全レベルの第1グループの入力部を選択的に駆動可能にしたことを特徴とする請求項6の装置。

【請求項9】 さらに複数のネットワークを有し、1つのネットワークのあるレベルのグループのうちの最後のグループからの出力が別のネットワークの全レベルの第1グループの入力部を選択的に駆動可能にしたことを特徴とする請求項6の装置。

10 【請求項10】 再構成可能なメモリプロセッサであって、

(a) 複数のメモリ素子と、

(b) それぞれ前記メモリ素子に対応した複数のプロセッサと、

(c) 前記メモリ素子の出力部を前記プロセッサの入力部に接続して、各プロセッサが前記メモリ素子の1つからの出力を入力として受け取るようにする第1セクタ手段と、

20 (d) 前記プロセッサの各々の出力部を前記プロセッサに対応したメモリ素子の入力部に接続すると共に、各メモリ素子の出力部を接続した第2セクタ手段と、

(e) 前記メモリ素子への入力として前記メモリ素子及びプロセッサの出力のうちの一方を選択するように前記第2セクタ手段を制御し、これにより複数のプロセッサが効果的に単一のプロセッサに減少され、かつその単一のプロセッサに使用可能なメモリ量をメモリ素子の数の因数により増大させるようにした手段と、を備えていることを特徴とする装置。

30 【請求項11】 前記第2セクタ手段は、それぞれ前記複数のメモリ素子に接続されている複数のマルチプレクサからなることを特徴とする請求項10の装置。

【請求項12】 前記制御手段がデコーダからなることを特徴とする請求項11の装置。

【請求項13】 複数のメモリ素子及びプロセッサで1グループが構成されており、前記グループに単一の第1セクタ手段及び単一のデコーダが設けられていることを特徴とする請求項12の装置。

【請求項14】 メモリチップレベルでのシステムエラーを検出する方法であって、

40 (a) チップに入ってくる多重ビットインターフェースでのパリティエラーを検出して、前記検出されたパリティエラーの各々の状態を保持する段階と、

(b) メモリアレイの行デコーダ回路のエラーを検出してその状態を保持する段階と、

(c) 単一ビットメモリエラーを検出して訂正し、2重のビットメモリエラーを検出してその状態を保持する段階と、を有することを特徴とする方法。

50 【請求項15】 さらに、1行のメモリ素子を細分割して、それぞれ複数列を有する訂正サブグループを形成する段階を有しており、それらの列が1つ置きに別々のエ

ラー検出訂正回路に接続されていることを特徴とする請求項14の方法。

【請求項16】 さらに、チップから前記エラー状態を読み取り、同時にエラー状態をクリアする段階を有していることを特徴とする請求項15の方法。

【請求項17】 さらに、保守のために単一ビットエラー状態と多重ビットエラー状態とを別々に維持する段階を有していることを特徴とする請求項15の方法。

【請求項18】 メモリチップでの処理において効果的な命令のデコード及び基本コードセグメントの実行を行う方法であって、

(a) 入力コマンド、アドレス及びデータ情報を別々のパイプラインに記憶する段階と、

(b) コマンド情報に従って第1の適当な時に前記情報をアクセスして、コマンドシーケンスによって定められた高レベルの機能を実行する段階と、

(c) 前記情報を演算して出力結果を発生し、それをデータパイプラインに再入力する段階と、

(d) コマンド情報に従って第2の適当な時にデータ及びアドレスパイプラインをアクセスして、この情報をメモリ素子に記憶する段階とを有することを特徴とする方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、メモリと計算を同一の集積回路上で結合して、データピン間に必要であった入力／出力部のほとんどをなくすことによってプロセッサとメモリとの間の命令／データ帯域幅を最大にするプロセッサインメモリ（PIM）チップに関するものである。

【0002】

【従来の技術】 並列SIMDシミュレーションワークベンチ(PASSWORK)についての研究から、多重命令多重データ(MIMD)ベクトル装置が、市販の単一命令多重データ(SIMD)装置の大域経路指定及びビットシリアル演算をほぼ全速でシミュレートできることがわかった。ハードウェアの集中／分散及びベクトルレジスタのコーナタニングが、「ビットシリアル直交変換命令を実行するための装置」と題する係属中のアイオブスト(Iobst)の米国特許出願番号第533,233号に開示されているベクトル装置でこの種の高性能SIMD計算をするのに重要である。

【0003】 ベクトル装置とSIMD装置とを直接的に比較すると、SIMD性能に対するその他の重要な限界は、メモリ帯域幅と、一部の算術、すなわちベクトル装置における全加算またはSIMD装置におけるプロセッサ間の計算に必要な多重論理演算とである。この研究の結果から、同一の共用メモリ装置でMIMD及びSIMDの両方の計算を可能にするには、SIMDを完全に新しい装置として構成するのではなく、従来装置にSIMDを盛り込む方法がよいことが暗示されている。

【0004】 プロセッサとメモリとを同一のチップに集積すれば、従来装置に基づいたSIMD性能をさらに高めることも可能である。すなわち、単一ビットプロセッサを標準形ランダムアクセスメモリ(RAM)集積回路(IC)の各列と対応させる新しい形式のメモリチップ(プロセッサインメモリ-PIMチップ)を設計すれば、SIMD性能を数倍も向上させることができるであろう。この性能向上は、電力、冷却及び／または空間要件をさほど増加させることなく達成できることに注目する必要がある。

【0005】 この基本構想は、メモリ内で直接的に計算することによって中央処理装置(CPU)とメモリとの間の非ノイマンボトルネックを破壊して、従来形計算環境からMIMD/SIMD混合計算環境へ自然に進化させることができる。この混合計算環境の応用は、今開発が始まったところである。

【0006】

【発明が解決しようとする課題】 このような事情に鑑みて本発明は、データピン間の入出力端子をほとんどをなくすことによってプロセッサとメモリとの間の命令／データ帯域幅を最大にするためメモリ部と演算部を同一の集積回路上に結合したPIMチップを提供するものである。

【0007】 本発明のさらなる目的によれば、アドレス及びデータバスを介してPIMチップのメモリバンクを制御するための方法が提供されている。

【0008】

【課題を解決するための手段】 上記目的を達成するため、本発明のPIMは、標準形メモリアレイに変更を加えてその1つの参照行で読み取られた全データが、その行の1ビットが1つのビットシリアルプロセッサに対応するようにして設けられた1組のビットシリアルプロセッサへ送られるようにしたプロセッサインメモリ(PIM)を構成する。プロセッサから計算された結果がメモリに書き込まれるとき、各プロセッサに1つずつ得られるすべての結果ビットも同時に書き込まれる。

【0009】 メモリはこのように並列にアクセスされるので、単一エラー訂正二重エラー検出(SECEDD)能力が付け加えられて、適正な作動を確保するか、あるいは不適当な作動であることを知らせる。性能が重要であるから、実行シーケンスは、メモリデータのロード及びストアに他のデータの処理をオーバーラップさせる。このパイプライン処理を実行するためにRレジスタが付け加えられている。

【0010】 PIMチップは2つのモードで作動できる。外部世界に対して、ICは通常の読み取り／書き込みメモリとして使用できる。計算プロセッサ及び追加制御回路によって能力を追加することによって、メモリアクセスサイクル(読み取りまたは書き込み)に代わってプロセッサで結果を計算することができる。

5

【0011】PIMチップを計算用を使用する場合（PIMモード）、チップピンから行デコーダにアドレスが与えられる。その結果、1行のデータがメモリから取り出される。データはエラー訂正されてから、クロックサイクルの終わり／次のクロックサイクルの始めでRレジスタ内にラッチされる。

【0012】次のクロックサイクルでは、プロセッサは外部コマンド／制御回線の制御を受けながらそのデータを計算シーケンスの一部として使用する。計算結果をプロセッサからメモリに記憶したい場合、メモリロードサイクルに代えて記憶サイクルにする。メモリアレイへ移動中の記憶データにエラー訂正チェックビットが加えられる。

【0013】多数のプロセッサ間、及び多数のプロセッサから1つまたは複数レベルのチップ外制御器へのデータの流れを容易にするためにプロセッサに追加される2つの通信ネットワークがある。これらのネットワークは、グローバル・オアネットワーク（GOR）及び並列プリフィクスネットワーク（PPN）として知られている。

【0014】チップが通常の書き込みに使用される時、まずデータがメモリから読み出され、エラー訂正が行われてから、Rレジスタに入れる前に書き込みデータと併合される。次に、変更データを入れたRレジスタの内容が、途中でエラー訂正論理を通してメモリへ送り戻される。これが必要であるのは、書き込みポートからチップに入ってくるビット数がメモリに書き込まれるデータ量よりも少ないためである。

【0015】この併合パスによって、適当なエラー訂正情報を書き込み中のワード用に再構成することができる。通常の読み取りに用いられた場合、1行のデータがメモリから取り出されて、エラー訂正をしてからRレジスタに入れられる。次のクロックサイクルで、チップから送り出すビットの適当なサブセットをアドレスビットが選択する。

【0016】通常の読み取り／書き込みモードでは、PIMチップバンクはアドレスバスに行及び列のアドレスを受け取り、データバスに読み取り／書き込みデータを受け取る。PIMモードでは、PIMチップバンクはアドレスバスに行アドレスを受け取り、データバスに実行すべきSIMD命令を受け取る。SIMD命令をアドレスバスの未使用の列アドレス回線で送ってもよい。このように、SIMD命令のシーケンスは、すべてのプロセッサ間にメモリ参照及び演算をオーバーラップして完全にパイプライン化することができる。

【0017】

【作用】チップには多数の単一ビット計算プロセッサが含まれ、それらはすべて並列駆動され、各チップ上のプロセッサの数は数個から数千まで渡っている。チップを合わせてメモリバンクのグループ化またはシステム化に

6

することによって、コンピュータ内の既存のメモリサブシステムをパーソナルコンピュータからスーパーコンピュータに変えることができる。

【0018】

【実施例】本発明の実施例を図面を参照して説明する。図1において、プロセスインメモリ（PIM）回路を説明する。回路の基本構成素子は、付加ローカルメモリ素子4を備えたビットシリアルプロセッサ2である。ローカルメモリ素子は、各クロックサイクル中にエラー訂正回路（ECC）論理を介して1ビットをビットシリアルプロセッサへ、またはそれから移動することができる。

【0019】このため、PIM構造のクロック速度はメモリアクセスにECC時間を加えた値に設定されている。あるいは、各クロックサイクル中に、やはりECC論理で処理した後、外部読み取り及び書き込みを実施できる。また、チップ上のプロセッサ素子間及びチップ間の通信経路を形成する論理が追加されている。

【0020】1つのビットシリアルプロセッサに対応したメモリは、1ビット幅のメモリ列で示されている。列が連結されてメモリアレイ6が形成されている。同様に1組のビットシリアルプロセッサが連結されており、通常は機能的にメモリアレイの下方に位置しているように見える。これは、メモリアレイへの単一行アドレスが、すべて並列になっているビットシリアルプロセッサの各々に1ビット与えることを意味している。

【0021】すべてのメモリアクセス、内部及び外部参照、及び読み取り書き込み動作は並行動作である。これは、PIM命令中、列アドレスビットは使用されないことを意味している。外部参照のための通常の列デコーダ及びセレクトを移動させることによってチップ構造、ECC処理及び結果的なタイミング変化を変えることができる。メモリアレイには、以下に詳細に説明するように余分のチェック列8が設けられている。

【0022】メモリアレイ6とプロセッサ2との間には、以下に詳細に説明する行デコードチェッカ12を備えたエラー検出訂正回路10が配置されている。

【0023】Rレジスタ14がエラー検出訂正回路10とプロセッサ2との間に配置されて、メモリデータのロード及びストアを他のデータの処理とオーバーラップさせるパイプライン化を行っている。

【0024】PIMチップは2つのモードで、すなわち通常の読み取り／書き込みメモリとして、または計算用として（PIMモード）作動できる。計算プロセッサ2と付加制御回線16によって能力を追加することによって、メモリアクセスサイクルの代わりにプロセッサで結果を計算できるようになる。

【0025】チップを計算用を使用する場合、チップピンから行デコーダ18にアドレスが与えられる。その結果、1行のデータがメモリから取り出される。データはエラー訂正されてから、クロックサイクルの終わり／次

のクロックサイクルの始めてRレジスタ内にラッチされる。次のクロックサイクルでは、プロセッサは外部制御及びコマンド回線16の制御を受けながらそのデータを計算シーケンスの一部として使用する。計算結果をプロセッサからメモリに記憶したい場合、メモリロードサイクルに代えて記憶サイクルにする。メモリアレイへ移動中にエラー訂正データが記憶データに加えられる。

【0026】チップを通常の書き込みに使用する場合、まずデータがメモリ4から読み出され、エラー訂正が行われてから、Rレジスタ14に入れる前に書き込みデータと併合される。次に、新しいデータを入れたRレジスタの内容が、途中でエラー訂正論理を通してメモリへ送り戻される。これが必要であるのは、書き込みポートからチップに入ってくるビット数がメモリに書き込まれるデータ量よりも少ないためである。この併合パスによって、適当なエラー訂正情報を書き込み中のワード用に再構成することができる。

【0027】通常の読み取りに用いられた場合、1行のデータがメモリから取り出されて、エラー訂正をしてからRレジスタに入れられる。次のクロックサイクルで、列アドレスビットはチップから送り出されるビットの適当なサブセットを読み取りセクタ22から選択する。

【0028】図示の実施例では、256のプロセッサが設けられており、これにSECCDEDチェックバイト列を加えると、メモリアレイは合計312列になる。各列の高さは2Kビットであると予想される。このため、メモリは $2048 \times 312 = 638,976$ (624K) ビットを有する。メモリアレイは物理的にこの大きさの構成する必要はなく、他のものにしても良好に動作する。

【0029】PIMチップ上の各プロセッサはビットシリアル計算ユニットである。すべてのプロセッサは同一であって、並行に制御される。すなわち、すべてのプロセッサはそれぞれ異なったデータについて同時に同一の演算を実行する。このように、プロセッサはSIMDコンピュータアーキテクチャを実施する。

【0030】次に図2を参照しながら、1つのビットシリアルプロセッサについて詳細に説明する。プロセッサには幾つかのマルチプレクサ24、26、27、28、30、31、32、33、34、36、37が設けられており、それらは、計算結果を条件に合わせて他のプロセッサまたはメモリへ伝播する手段を含む固定機能算術論理ユニット(ALU)38に出力を送る。

【0031】ALU38はA、B及びCと呼ばれる3つの入力信号を受けて、これら3つの入力の3つの固定機能結果を計算する。その結果は、合計($A+B+C$)、繰り上げ($A \cdot B + A \cdot C + B \cdot C$)及びブasting比較($C+A \cdot B$)である。マルチプレクサの能力を利用して、全組の論理演算を繰り上げ機能から実行することができる。例えば、C入力を遮断($C=0$)することに

よって、A及びBの論理積が計算され、C入力を導通($C=1$)させることによってA及びBの論理和が計算される。

【0032】いくつかのマルチプレクサがプロセッサ内のデータ経路及び機能を選択する。マルチプレクサを作動させるデータソースは、メモリから送られるか、内部通信ネットワークを介して他のプロセッサから送られるか、内部的に発生して保管されている結果から送られる。

10 【0033】ALUのA、B及びC入力を送る3つの主要マルチプレクサ24、26、28がある。各マルチプレクサはそれぞれの制御/コマンド回線で制御されている。図面において、制御回線がFnで示されており、ここでnは1から20までの数字である。すべての制御回線はチップから出ている。マルチプレクサ24、26、28の各々は3つの独立した制御回線で駆動される。そのうちの2つの回線はデコードされて4つの入力の内の1つを選択するのに対して、第3の回線は選択信号の状態を反転させる。

20 【0034】第1マルチプレクサ24は、制御回線の制御を受けながら、直前のクロックサイクルから得られたマルチプレクサ24の先行出力(この状態はマルチプレクサ24に対応したフリップフロップ40で保管されている)、メモリから読み出されたデータ、別の制御/コマンド回線で駆動される別のマルチプレクサによっていずれかが選択されるALUからの合計または繰り上げ結果、及び論理ゼロから選択することができる。これらの信号はいずれも各クロックサイクル時にALUのA入力へ、おそらく反転させて送ることができる。

30 【0035】第2マルチプレクサ26は、様々な通信経路から選択するか、前に計算されている一部の結果を戻す二次レベルマルチプレクサ27から第1入力を受ける以外、第1マルチプレクサ24と同じデータ入力を受ける。制御回線は、第1マルチプレクサに接続された制御回線と同一機能を果たすが、それらから独立している。第1マルチプレクサの場合と同様に、ALUに送られるデータは必要に応じて反転させることができる。

40 【0036】第3マルチプレクサ28は、直前のクロックサイクルから得られたマルチプレクサの先行出力(この状態は第3マルチプレクサ28に結合したフリップフロップ42に保持されている)、第2マルチプレクサ26へ送られる通信マルチプレクサ27からのデータと同じもの、別の制御/コマンド回線で駆動される別のマルチプレクサによっていずれかが選択されるALUからの合計または繰り上げの2信号の比較結果、及び論理ゼロから選択することができる。選択されたデータは、3つの独立した制御回線の制御を受けながらおそらく反転させてALUへ送られる。

50 【0037】いずれのSIMD装置も、一部のプロセッサが特定の演算を実施しないようにする機構を必要とす

る。PIMに選択されたこの機構は、条件付き記憶である。すなわち、一部のプロセッサがコマンドを実行することを禁止するのではなく、すべてのプロセッサにコマンドを実行させるが、計算結果を記憶しないというものである。

【0038】この形式の条件制御を実施するため、3つのフリップフロップ35がマルチプレクサ31、33、36及び37と共にプロセッサに加えられている。いずれのサイクルにおいても、マルチプレクサは3つの内のいずれを選択することもでき、あるいは論理ゼロを選択することもできる。前述のマルチプレクサと同様に、選択された入力の状態を反転させることができる。このため、例えば入力として論理ゼロを選択すると、反転信号／コマンドを活性化させることによって出力を論理1にすることができる。

【0039】SIMD命令シーケンスが実行されると、メモリからの古いデータがAマルチプレクサに対応したフリップフロップにロードされて、計算結果がALUからBマルチプレクサへ送られる。マルチプレクサ36から入力を受けるマルチプレクサ33が論理1を出力している場合、Bデータがメモリ記憶経路を進むことが阻止されるか、Aマルチプレクサからのデータの進路が阻止される。

【0040】データは、一般的にメモリからマルチプレクサ26を介して、またはALUから計算結果としてマルチプレクサ26または28を介してロードされたデータから記憶許可フリップフロップ35にロードされる。コマンド回線が別のマルチプレクサ34を介していずれかの結果を選択し、さらなるコマンド回線が（もしあるならば）どの記憶許可ビット35にロードするかを選択する。

【0041】データは、各プロセッサから、PIMチップ上またはそれ以外のプロセッサ間の通信を行うネットワークへ送られる。ネットワークには、グローバル・オアネットワーク（GOR）及び並列プリフィクスネットワーク（PPN）の2種類がある。GORは多数から個、または個から多数への通信に用いられるのに対して、PPNは多数から多数への通信に用いられる。

【0042】GORへ送られるデータは、記憶許可ビット35の1つで進路が阻止される。これにより、特定のプロセッサの記憶許可ビットを論理1にする一方、その他のプロセッサの記憶許可ビットを論理ゼロにすることによってその特定プロセッサがGORネットワークを駆動することができる。

【0043】あるいは、チップ上のすべてのプロセッサがGORネットワークを駆動して、すべてのプロセッサのグローバル・オアを個々のプロセッサまたはチップ外の高レベルの制御器へ送り戻すこともできる。チップ上のすべてのプロセッサ間のグローバル・オアは多重レベルORゲート49を介して実行される。

【0044】GOR及びPPNの両ネットワークからのデータは、個別のコマンド回線で制御される別のマルチプレクサ27によって選択される。このデータは、第2及び第3マルチプレクサ26、28のいずれか（または両方）で選択できる。

【0045】次に図3を参照しながら、並列プリフィクスネットワークについて説明する。このネットワークの名前は、走査または平行プリフィクスと呼ばれる数学機能に由来している。図3のネットワークは、多くの並列を可能にして関係オペレータ間の並列プリフィクスを高速化できるようにしてこの機能を実行する。

【0046】加算についてのプリフィクス演算は走査と呼ばれ、以下のように定義される。

$$【0047】X_i = X_{i-1} + Y_i$$

但し $i = 1 \sim n$ 、 $X_0 = 1$

または

$$X_1 = Y_1$$

$$X_2 = X_1 + Y_2$$

$$X_3 = X_2 + Y_3$$

$$X_4 = X_3 + Y_4$$

演算の連鎖に注目されたい。このように説明した場合、各結果はすべての先行結果の影響を受ける。しかし、等式は次のように展開できる。

$$【0048】X_1 = Y_1$$

$$X_2 = Y_1 + Y_2$$

$$X_3 = Y_1 + Y_2 + Y_3$$

$$X_4 = Y_1 + Y_2 + Y_3 + Y_4$$

各プロセッサは単一データ項 $Y_1 \sim Y_n$ でスタートする。PPNは、 Y_2 のコピーを保持しているプロセッサがそのデータを Y_1 を保持しているプロセッサに送ることができるようにすると同時に、 Y_4 を保持しているプロセッサがそのデータを Y_3 を保持しているプロセッサに送る等ができるようにする。各プロセッサは、データについて必要な演算（この例では加算）を他の同様な演算と並行して実行して、さらなる計算に使用できる途中結果を得ることにより、すべてのプロセッサで結果が得られる、すなわちプロセッサ1には X_1 が、プロセッサ2には X_2 が得られる。

【0049】このネットワークをハードウェアで実行してそれを一般のプロセッサ通信に利用することによって、2つの利点が得られる。このネットワークは、第1に、直列に連続して実行せざるを得ない幾つかの機能を並行して実行でき、第2に、シリコンで非常に効果的にチップにわずかな経路空間をとるだけで並列機能が達成されるネットワークはプロセッサ間ですべての対数レベルで実行される。一次レベルではプロセッサはデータを左側のプロセッサに送ることができる一方、右側のプロセッサからデータを受け取る。次のレベルでは、特定のプロセッサがその左側に隣接した2つのプロセッサにデータを送る。さらに次のレベルでは、データを受け取

るプロセッサの数が二倍になる一方、データを送るプロセッサの数は半分に減る。すべてのプロセッサはすべてのレベルからデータを受け取る。外部で実行中の実行プログラムによって状態が制御されている制御回線が必要なレベルを選択する。すべてのプロセッサは同一レベルを選択する。

【0050】PPNの基本実行からの幾つかの拡張が設けられている。このため、あるレベルを完成させるために必要な接続が実行される。すなわち例えばレベル0では、偶数番号のプロセッサがそれに左側のプロセッサにデータを送ることができるが、これはPPN機能では必要でない。また、データ移動を逆方向に、すなわち右方向へ実行する別のレベル0がPPNネットワークに加えられる。さらに、右方向及び左方向データ移動接続部の端部に、通信を拡張モードまたは循環モードで実行できるようにするマルチプレクサ46、48が付け加えられている。循環モードでは、チップの最後のプロセッサが最初のプロセッサを駆動する（反対方向のデータ移動の場合、最初のプロセッサが最後のプロセッサを駆動する）。拡張モードでは、端部プロセッサはチップ外からデータを受け取る。これによって、1チップ以上の通信ネットワークを構築することができる。

【0051】プロセッサの数及び実際の最大チップ寸法によって定められる限界のため、各プロセッサに使用できるメモリ量は限られている。また、使用可能な数のプロセッサを完全に利用することができないようにするプログラム及びアルゴリズムがある。両問題を同時に解決する試みは列減少と呼ばれ、図4を参照しながら以下に説明する。

【0052】プロセッサはグループ化されるため、以前は各プロセッサの専有であったメモリがグループ間で共用される。付加アドレス回線として機能する付加制御回線が要求されたデータを特定のメモリ列からグループ内のすべてのプロセッサへ送る。このため、グループ内の各プロセッサは同一データで計算する（グループの一部であるかどうかに関係なく、すべてのプロセッサは同一機能を実行することを覚えておかれたい）。データを記憶したい場合、記憶するデータのアドレスに対応したプロセッサを使用可能にして、新しく計算された結果をメモリへ送る一方、グループ内で記憶アドレスに対応していないプロセッサは、前に記憶アドレスから取り出された古いデータを複写する。

【0053】すなわち、複数のメモリ素子50、52、54、56にはそれぞれプロセッサ58、60、62、64が対応して設けられている。第1セクタ66がメモリ素子の出力部をプロセッサの入力部に接続するため、各プロセッサはメモリ素子の1つからの出力を入力として受け取る。複数のマルチプレクサ68、70、72、74が各プロセッサの出力部をそれに対応したメモリ素子の入力部に接続している。各メモリ素子の出力部

はフィードバック回線76を介して対応のマルチプレクサにも接続している。デコーダ78がマルチプレクサ68、70、72、74を制御して、メモリ素子及びプロセッサの出力の1つをメモリ素子への入力として選択する。このため、複数のプロセッサは効果的に単一のプロセッサに減じられ、その単一プロセッサに使用できるメモリ量は、メモリ素子の数の倍数に増加する。

【0054】複数のメモリ素子及びプロセッサを配列して、単一のセクタ及び単一のデコーダを備えた1つのグループにすることができる。

【0055】上記の実行は、すべてのメモリを処理グループから1つのプロセッサへ送り、そのプロセッサからの結果を正しい記憶アドレスへ送る論理に置き換えることもできる。この実行は、機能的に正確であるが、論理経路に余分なタイミングキューを持ち込み、上記のデータの条件記憶の実行が非常に複雑になる。

【0056】通常の外部エラー訂正に代わって、メモリ（すべての外部読み取り装置を含む）から読み取られているすべてのデータを訂正し、メモリ（やはり外部読み取り装置を含む）へ書き込まれるすべてのデータのためのチェックバイトを発生し、メモリ（やはり外部書き込み装置を含む）へ書き込まれるすべてのデータのためのチェックバイトを発生する1組の内部SECDEDブロックが設けられている。SECDEDは、32データビット及び7チェックビットの39ビットグループの反復セットとして実行される。データビットには対応のビットシリアルプロセッサが設けられているが、チェックビットにはない。対になった39ビットグループはそれらのビットがインターリーブされている。このため、78ビットグループ（ $78 = 2(32 + 7)$ ）では、偶数番号のビットが1つのSECDEDグループに対応しており、奇数番号のビットが別のものに対応している。これは、隣接した短絡ビット線のようなエラーが2ビットの回復不能のエラーとしてではなく、2つの単一回復可能エラーとして現れることを意味する。トレードオフとして、インターリーブド72ビットグループを考えることができる。メモリグループは144列（ $144 = 2(64 + 8)$ ）になる。合計が312列ではなく288列であって、（提案されている4グループではなく）2メモリグループになる。

【0057】その他のオンチップエラー検出論理もある。受信データ及びアドレスのパリティは、SIMDコマンドのパリティと同様に受信後に別々に検査される。チップからの読み取りデータのパリティはデータと一緒に送られる。アクセスされた行パリティチェックもある。受信アドレスの行部分のパリティが、実際にアクセスされた行のパリティを内容として含む特定のメモリ列の内容と比較される。いずれかのパリティによって検出されたエラーまたはSECDED不良はチップ状態レジスタに設定される。チップ状態は通常読み取り経路か

ら確認されるか、チップ保守ポートからアクセスすることができる。

【0058】外部読み取り及び書き込みタイミングはエラー訂正論理の影響を受ける。読み取り動作時には、データはメモリから読み取られ、エラー訂正された後、Rレジスタに入れられる。最初の2つのアドレスビットがこのレジスタに入る途中に分解される。第2サイクルにおいてアドレッシング選択が完了して、データがその部分から送り出される。アドレッシング及びデータ経路によって、インターリーブDSECDDEDグループの64データ列がチップ上及びチップ外の1つのデータビットを駆動するようになっている。

【0059】外部書き込みの場合、読み取りアドレスのワードが読み取られ、エラー訂正された後、4書き込みビットと併合されてRレジスタに入る。次のクロックサイクルで、レジスタに保持されているデータからチェックビットが発生し、合計312ビットが書き込まれる。外部アドレスを第2メモリサイクルから有効に保つレジスタが設けられているので、チップピンのデータ及びアドレスは1サイクル期間中だけ有効であればよい。

【0060】上記の最後の2文節は、PIMチップが外部世界に対して同期インターフェースを示すことを指摘している。読み取りの場合、データは読み取り動作を開始するクロックから第2クロックエッジ後に有効となる。少なくともチップレベルにおいて、新しい読み取りサイクルがすべてのクロックで開始されるが、データエラーがある場合には訂正データをメモリに書き込むことが望ましく、これには別に1つのクロックサイクルが必要になる。書き込みの場合、チップは2クロックサイクルに渡って作動するが、データは両サイクルに渡って有効である必要はない。もちろん、PIMチップクロックがコンピュータシステムの他のものと同じクロック速度であると言っているわけではない。

【0061】また、PIMチップには幾つかのエラー検出機構があり、以下のものが含まれる。

【0062】データパリティの検出及び発生：第5ビットは読み取り及び書き込みの両方において4ビットデータインターフェースを伴う。

【0063】アドレスパリティ：パリティビットは、受信アドレス毎に外部読み取り書き込み用か、PIMモード参照用かが検査される。

【0064】コマンドパリティ：SIMDコマンド毎にパリティビットが検査される。

【0065】行パリティ：参照された行のパリティを内容に含む特別の列がメモリアレイに加えられる。このビットは、受信行アドレスのパリティと比較される。ここでは列減少モードと違いがない。

【0066】これらのエラーのすべては、SECDDED論理で検出された単一ビット及び多重ビットエラーと一緒にPIM状態フリップフロップに入れられる。これら

は通常のメモリアクセス回線を介して読み取られるか、チップ保全ポートを介して読み取られる。

【0067】保全ポートはJTAG/IEEE 1149.1にすることができる。チップ状態に加えて、一部のチップ試験情報をこのポートからアクセスできる。

【0068】一部のデータ経路を制御したり、他の方法では試験が非常に難しい（あるいは不可能である）いくつかの診断的機能を実施するための様々なビットがチップに埋め込まれている。チェックバイト発生を停止させるための制御ビットが設けられている。これによってSECDDED論理を検査することができる。なすべきことは、書き込みチェックバイトを全ゼロデータワードで発生する値と同じ値にすることである。制御ビットは、行パリティ論理内で比較を反転させることもできる。その時、いずれのPIM参照も行パリティエラー状態ビットを設定しなければならない。他のビットがPPNデータ経路指定を行う。

【0069】要約すると、メモリチップレベルでのシステムエラーを検出する方法は、チップに入ってくる多重ビットインターフェースでのパリティエラーを検出する段階と、検出されたパリティエラーの各々の状態を保持する段階とを有している。次に、メモリアレイ行デコーダ回路のエラーが検出されて、そのエラー状態が保持される。単一ビットメモリエラーが検出、訂正され、2ビットメモリエラーが検出されて、その状態が保持される。

【0070】メモリ素子の行は訂正サブグループに細分割されて、その各サブグループは複数列を有しており、それらの列は1つ置きに別々のエラー検出訂正回路に接続されている。次にチップからのエラー状態が読み取られ、同時にクリアされる。単一ビットエラー状態及び多重ビットエラー状態は保守用に個別に維持される。

【0071】PIMモード実行は、メモリ参照が読み取りか書き込みかを判断するためにR/W回線が用いられる点で、通常の読み取り/書き込み制御に非常によく似ている。PIM読み取りモードでは、アドレス回線が制御用に用いられ、データ回線が状態/制御情報をCPUに戻すために用いられる（1ビット/PIMデータ回線）。PIM書き込みモードでは、データ回線がPIM制御に用いられ、アドレス回線がプロセッサ間の行選択の特定化に用いられる。

【0072】メモリチップでの処理における効果的な命令デコード及び基本コードセグメントの実行を行う方法も考えられる。この方法によれば、入力されるコマンド、アドレス及びデータ情報は別々のパイプラインに記憶される。これらの情報はコマンド情報に従って第1の適当な時にアクセスされて、コマンドシーケンスによって定められた高レベルの機能を実行する。情報は算術論理ユニットによって演算されて、それによって得られた出力結果がデータパイプラインに再入力される。データ

及びアドレスパイプラインはコマンド情報に従って第2の適当な時にアクセスされて、この情報がメモリに記憶される。

【0073】以上に本発明を好適な実施例について説明してきたが、発明の概念から逸脱しない範囲内において様々な変更及び変化を加えることができることは当業者には明らかであろう。

【0074】

【発明の効果】メモリチップ内の処理は、同一集積回路上でメモリ部と演算部とを結合させているので、データ 10 ピン間の入出力端子がほとんどなくなり、プロセッサとメモリとの間の命令/データ帯域幅を最大にすることができる。

【0075】チップには、並行に駆動される多数の単一ビット計算プロセッサが含まれているため単一命令多重データの計算を迅速に処理することができ、さらにエラー訂正論理もチップに組み込まれているためメモリデー

タ内のエラーを発生時に検出して訂正できる。したがって、本発明は高性能なPIM装置を構成する。

【図面の簡単な説明】

【図1】本発明によるPIMチップのブロック図である。

【図2】図1のPIMチップのビットシリアルプロセッサの概略図である。

【図3】図1のPIMチップのグローバル・オア/並列プリフィクスネットワークを説明する概要図である。

【図4】メモリの列減少を行うための再構成可能メモリプロセッサのブロック図である。

【符号の説明】

2 プロセッサ

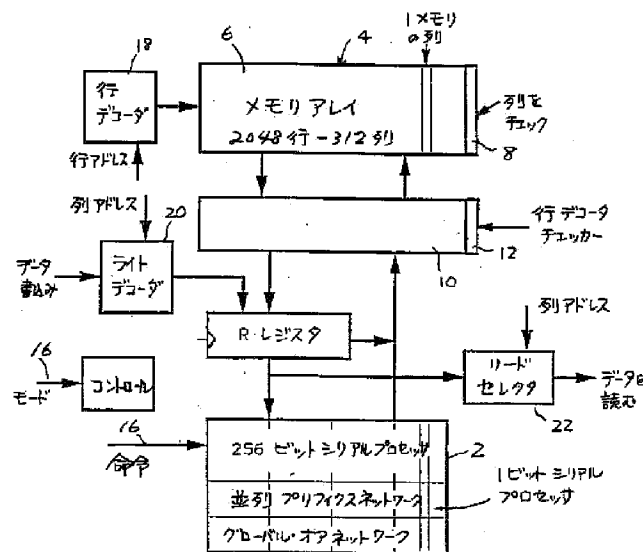
4 メモリアレイ

24, 26, 27, 28, 30, 31, 32, 33, 34, 36, 37 マルチプレクサ

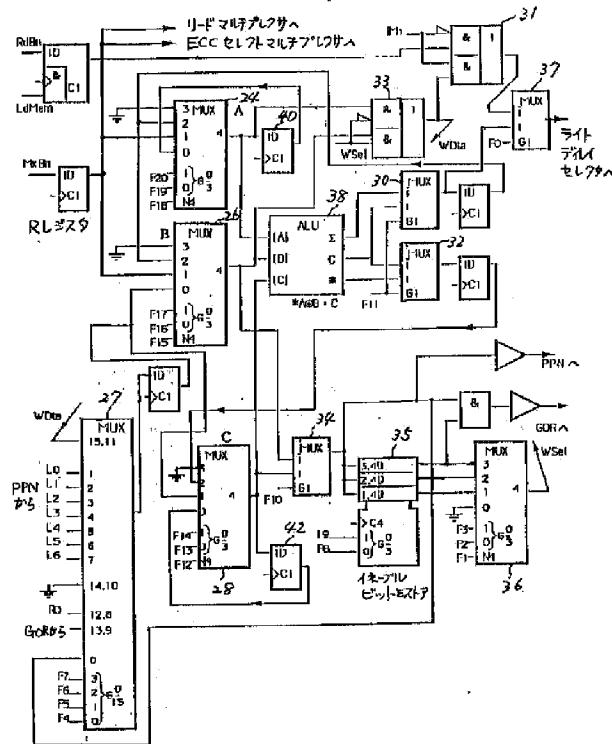
35 フリップフロップ

38 算術論理ユニット

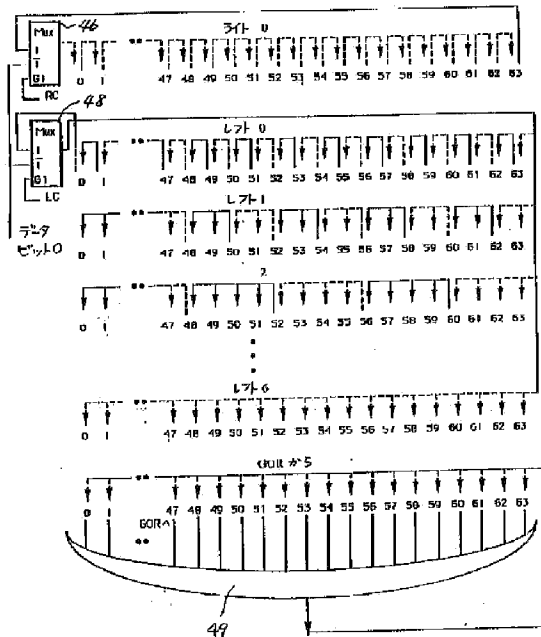
【図1】



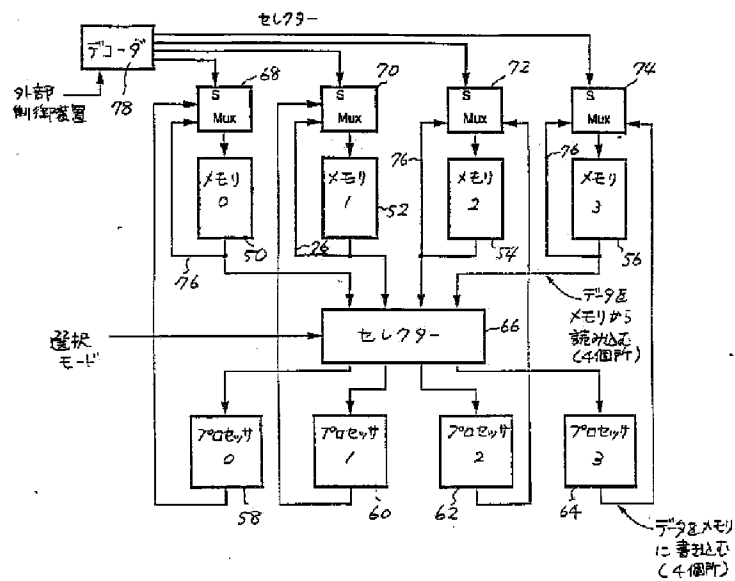
【図2】



【図3】



【図4】



フロントページの続き

(51)Int. Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 1 1 C 29/00	3 0 2	8526-5L		

(71)出願人	591176982	(72)発明者	ケネス ダヴリユー、 アイオブスト
	ケネス アール、 ウォルグレン		アメリカ合衆国 メリーランド 20904
	KENNETH R. WALLGREN		シルバースプリング ウインドミル コー
	アメリカ合衆国 メリーランド 21045		ト 18
	コロンビア タリスマン レイン 7225	(72)発明者	デビッド アール、 レスニツク
			アメリカ合衆国 ウィスコンシン 54703
			オウ クレア ブラック アベニユー
			2006
		(72)発明者	ケネス アール、 ウォルグレン
			アメリカ合衆国 メリーランド 21045
			コロンビア タリスマン レイン 7225